Book Section:
Chambers, Marcus, Thornton, Michael Alan orcid.org/0000-0002-4470-809X and McCrorie, Roderick (2018) Continuous Time Modelling Based on an Exact Discrete Time Representation. In: van Montfort, Kees, Oud, Johan and Voelkle, Manuel, (eds.) Continuous Time Modeling in the Behavioral and Related Sciences. Springer , 317 - 357.

Forthcoming in: K. van Montfort, J. Oud and M. Voelkle (Eds.), *Continuous Time Modeling in the Behavioral and Related Sciences*, Springer.

Continuous Time Modelling Based on an Exact Discrete Time Representation

Marcus J. Chambers
University of Essex

J. Roderick McCrorie
University of St Andrews

and

Michael A. Thornton
University of York

October 2017

Abstract

This chapter provides a survey of methods of continuous time modelling based on an exact discrete time representation. It begins by highlighting the techniques involved with the derivation of an exact discrete time representation of an underlying continuous time model, providing specific details for a second-order linear system of stochastic differential equations. Issues of parameter identification, Granger causality, nonstationarity, and mixed frequency data are addressed, all being important considerations in applications in economics and other disciplines. Although the focus is on Gaussian estimation of the exact discrete time model, alternative time domain (state space) and frequency domain approaches are also discussed. Computational issues are explored and two new empirical applications are included along with a discussion of applications in the field of macroeconometric modelling.

Keywords. Continuous time; exact discrete time representation; stochastic differential equation; Gaussian estimation; identification; Granger causality; nonstationarity; mixed frequency data; computation; macroeconometric modelling.

Acknowledgements: We thank an Editor and two anonymous referees for helpful comments that have led to improvements in this paper, and the Scottish Institute for Research in Economics for arranging facilities in the School of Economics, University of Edinburgh, for the authors to meet to work on this chapter. The first author also thanks the Economic and Social Research Council for financial support under grant number ES/M01147X/1.

Address for Correspondence: Marcus J. Chambers, Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, England.
Tel: +44 1206 872756; e-mail: mchamb@essex.ac.uk.

1. Introduction

Time series modelling in the social sciences often involves data that are generated in finer time intervals than the sampling interval pertaining to the available data. In economics, for example, macroeconomic time series represent the aggregation of a large number of decisions made by microeconomic agents within the chosen sampling interval. Even today, however, the issues in macroeconomics addressed using aggregate time series data almost never tie models or conclusions to parameters governing the pre-aggregated behaviour of economic agents. Instead, at best, agents' preferences are modelled through a so-called representative agent.

On the other hand, when estimating and making inferences about parameters of interest, econometrics has tended to embody and adapt developments in the statistical analysis of time series. Notably, its response to Box-Jenkins models, which as 'black-box' forecasting models in the 1960s and 1970s outperformed structural econometric models that incorporated restrictions based on economic theory, was to create a unit root/co-integration paradigm that embodied the best features of both approaches. Even so, the one pervasive characteristic of time series econometrics has been its use of linear-in-variables discrete time series models, such as autoregressive (AR) or autoregressive moving average (ARMA) models and their vector counterparts, as the basis of model specification.

One aim of this chapter is to draw attention to a modelling issue that still perhaps does not take on the importance it deserves: the potential incompatibility of using such linear time series models, if naïvely specified, in a context where the data are generated in finer time intervals than the interval pertaining to the available data. This is because linear discrete time models are not time-invariant, meaning that, on a strict interpretation, parameter estimates are tied only to a particular sampling frequency. Such discrete time models therefore do not readily admit an economic interpretation in the absence of a treatment of temporal aggregation bias. One potential remedy to this problem is to formulate a structural model in continuous time with the property that equidistant data generated from its solution satisfy a linear discrete time model. The essence of the method relies on the derivation of a system of stochastic difference equations that satisfy exactly a linear stochastic differential equation system with constant coefficients. Such a discrete time model is called an *exact discrete time model* and, through it, the structural AR or ARMA specification can be embodied in statistical inference independently of the sampling interval.[1]

The approach based on an exact discrete time model has been historically associated with A.R. (Rex) Bergstrom[2] who, perhaps more than any other econometrician, presaged the advent of continuous time models in econometrics and finance; see, for example, Bergstrom (1966, 1983), although it was Peter C.B. Phillips (1972) who provided the first implementation of the methods discussed in this chapter.[3] There are, however, some costs in following this approach, notably that in multivariate models, identifying the parameters of the

[1] McCrorie (2009) lists a number of contributions that use an exact discrete time model.

[2] Rex Bergstrom spent over twenty years of his academic career at the University of Essex and had both direct and indirect influences on the current authors. He taught both Marcus Chambers and Roderick McCrorie at the Masters level and supervised the PhD thesis of Chambers (1990). Chambers, in turn, was the PhD supervisor of McCrorie (1996) and Michael Thornton (2009).

[3] This paper was based on Phillips's M.A. dissertation supervised by Bergstrom at the University of Auckland in 1969. It represented the first of many contributions by Phillips on continuous time econometrics; Yu (2014) provides a survey of this work.

structural continuous time model on the basis of discrete time data is considerably more challenging than identifying the parameters of an (albeit time-varying) discrete time model using the same data (see section 2.2 below). This chapter discusses the development of, and issues arising in, the formulation of structural continuous time models and the estimation of their parameters using an exact discrete time model, in a way that we hope will facilitate future applications in interdisciplinary areas.

Throughout this chapter we focus mainly on continuous time models specified as systems of linear stochastic differential equations, although in Section 4 we also briefly discuss non-linear systems in the macroeconomic modelling literature that have antecedents in linear-in-variables approaches. Recent developments have enabled non-linear systems to be estimated directly; see Wymer (1997, 2012) for details. Discussion of non-linear diffusion-type models in finance, which appear in the survey by Aït-Sahalia (2007), are outside the scope of this chapter.[4] See Aït-Sahalia and Jacod (2014) for a comprehensive treatment of this topic.

The advantages of formulating econometric models in continuous time, over and above the issue of embodying an ARMA-type specification independently of the sampling frequency, were discussed by Bergstrom (1990, 1996), *inter alios*. Specifically, continuous time models can take account of the interaction among variables during the observation interval; they permit a more accurate representation of the partial adjustment processes in dynamic disequilibrium models, as discussed in section 4.3 below; they allow a proper distinction to be made in estimation between stock variables (measured at points in time) and flow variables (measured as integrals of a rate of flow over the observation period); and they can be used to generate forecasts of the (unobservable) continuous time paths of the variables.

In view of the backgrounds and expertise of the authors this chapter is written from the viewpoint of economics and, more specifically, econometrics. It therefore mostly neglects the treatment of the estimation of continuous time models in other areas of the social sciences and science more generally, such as engineering. Material relevant to other disciplines can be found in other contributions to this volume. The plan of this chapter is as follows. Section 2 is broadly concerned with continuous time methods in econometrics and contains seven sub-sections. The first lays the groundwork for subsequent sections and explains how an exact discrete time model corresponding to a linear continuous time system can be obtained, and provides a worked example for a second-order differential equation system. Section 2.2 deals with the fundamental problem of identification of the parameters of a continuous time system from discrete time data, and section 2.3 discusses how the process of temporal aggregation can distort inferences relating to Granger causality. Section 2.4 explores various issues of nonstationarity that are important when analysing economic and financial time series, while section 2.5 summarises recent work that enables the information contained in observations made at different sampling frequencies to be used in the estimation of a continuous time system. The remaining two sub-sections deal with Gaussian estimation as well as alternative (frequency domain) methods.

The final sections of the chapter have a more practical aim. Section 3 is devoted to computational issues and reports the results of a small simulation exercise (the code for

[4] Most non-linear models are not directly amenable to the derivation of exact discrete time representations and typically result in transition densities that have no closed-form solution. See, however, Phillips and Yu (2009), Fergusson and Platen (2015) and Thornton and Chambers (2016), for examples where a closed form density is apposite.

which is contained in the appendix) while section 4 is concerned with empirical applications. Sections 4.1 and 4.2 contain new applications to consumer prices and inflation, and to oil prices and the macroeconomy, respectively, while section 4.3 discusses applications of the continuous time methodology in the arena of macroeconometric modelling. Section 5 contains some concluding comments.

2. Continuous time models in econometrics

2.1. Linear continuous time systems and exact discrete time models

We will be concerned with an $n \times 1$ vector of variables, denoted $x(t)$, whose dynamic evolution is determined by a stochastic differential equation system in continuous time. Bergstrom (1983, 1984) provided a rigorous foundation for the specification of such systems and pioneered the development of the exact discrete time approach for first- and second-order systems, subsequently extended by Chambers (1999) to systems of order greater than two. A higher-order system is specified as

$$d[D^{p-1}x(t)] = [A_{p-1}D^{p-1}x(t) + \ldots + A_1 Dx(t) + A_0 x(t)]dt + \zeta(dt), \quad t > 0, \qquad (1)$$

where $A_{p-1}, \ldots, A_0$ are $n \times n$ parameter matrices, D denotes the mean square differential operator satisfying

$$\lim_{\delta \to 0} E \left| \frac{x_i(t+\delta) - x_i(t)}{\delta} - Dx_i(t) \right|^2 = 0, \quad i = 1, \ldots, n,$$

$x(0), \ldots, D^{p-1}x(0)$ are a set of initial conditions,[5] and $\zeta(dt)$ is an $n \times 1$ vector of random measures with $E[\zeta(dt)] = 0$, $E[\zeta(dt)\zeta(dt)'] = \Sigma dt$ (Σ being an $n \times n$ symmetric positive definite matrix), and $E[\zeta(\Delta_1)\zeta(\Delta_2)'] = 0$ for any disjoint intervals, Δ_1 and Δ_2, on the real line $-\infty < t < \infty$.[6] Under these assumptions the random measure vector $\zeta(dt)$ is similar to vector white noise and the system (1) can be regarded as a continuous time autoregressive system of order p, which we shall denote CAR(p). The system could be extended to include a deterministic linear trend function with the addition of a term of the form $[\gamma_0 + \gamma_1 t]dt$ on the right-hand-side of (1), where γ_0 and γ_1 are $n \times 1$ vectors of unknown parameters, or to include exogenous variables, but to do so would result in additional complexity that we wish to avoid here. The system (1) is interpreted as meaning that $x(t)$ satisfies the stochastic integral equation

$$D^{p-1}x(t) - D^{p-1}x(0) = \int_0^t [A_{p-1}D^{p-1}x(r) + \ldots + A_1 Dx(r) + A_0 x(r)]dr + \int_0^t \zeta(dr)$$

for all $t > 0$; see Bergstrom (1983) for further details.

---

[5]The initial conditions are usually assumed to be fixed which imparts a type of nonstationarity on an otherwise stable system. This is a different type of nonstationarity to that which has dominated the econometrics literature in recent years and which we discuss in section 2.4.

[6]The use of a vector of random measures to specify the disturbance vector in a continuous time model in the econometrics literature is due to Bergstrom (1983) who built on the work of Rozanov (1967). A common alternative is to replace $\zeta(dt)$ with $\Sigma^{1/2}dW(t)$ where $dW(t)$ denotes the increment in a vector of Wiener processes and $\Sigma^{1/2}(\Sigma^{1/2})' = \Sigma$. Note, though, that the latter specification imposes Gaussianity on the system whereas the distribution of $\zeta(dt)$ is unspecified beyond its first two moments.

The objective is to estimate the elements of the matrices $A_{p-1}, \ldots, A_0$ and Σ from a sample of data observed at discrete points in time i.e. not observed continuously. The elements of these matrices will often be known functions of an underlying vector of structural parameters although we avoid emphasising such dependencies here for reasons of notational simplicity.[7] The exact representation approach derives the law of motion for the observations that is consistent with their having been generated by the stochastic differential equation system (1). The nature of the observations themselves depends on the form of variables that comprise the vector $x(t)$. In the most general (mixed sample) case the vector $x(t)$ can be partitioned into an $n^s \times 1$ subvector of stock variables (x^s) and an $n^f \times 1$ subvector of flow variables (x^f), where $n^s + n^f = n$, so that

$$x(t) = \begin{pmatrix} x^s(t) \\ x^f(t) \end{pmatrix}.$$

Stock variables are assumed to be observable at equally spaced discrete points in time of length h, resulting in the sequence

$$\{x^s_{th} = x^s(th)\}_{t=0}^{T} = \{x^s_0, x^s_h, \ldots, x^s_{Th}\},$$

while flow variables are observable as an integral of the underlying rate of flow over the sampling interval of length h, yielding the sequence

$$\left\{x^f_{th} = \frac{1}{h}\int_{th-h}^{th} x^f(r)dr\right\}_{t=1}^{T} = \left\{\frac{1}{h}\int_{0}^{h} x^f(r)dr, , \ldots, \frac{1}{h}\int_{Th-h}^{Th} x^f(r)dr\right\}.$$

Examples of stock variables in economics include the money stock, exchange rates, interest rates and other asset prices, all of which are observable (at least in principle) at points in time. Examples of flow variables include consumers' expenditure, income, exports, imports, and cumulative rainfall in Brazil, each of which is measured as the accumulation of a rate of flow over a time interval (corresponding with the sampling interval). Although we assume that the observations are equally spaced it is possible to extend the setup to allow for irregularly spaced observations. This can be achieved by introducing an index $i = 1, \ldots, N$, where N denotes sample size, and to denote the sampling intervals by $h_i = t_i - t_{i-1}$. For notational convenience, however, we shall assume that the observations are equally spaced. Also, for the purposes of clarity, we will, for the time being, assume that $x(t) = x^s(t)$ so that all n variables are of the stock variety. The consequences of relaxing this assumption will be discussed in due course.

The first step in deriving an exact discrete time representation is to write the model in a suitable state space form. In order to do this we can define the $np \times 1$ state vector

$$y(t) = [x(t)', Dx(t)', \ldots, D^{p-1}x(t)']',$$

which satisfies the first-order stochastic differential equation system

$$dy(t) = Ay(t)dt + \phi(dt), \quad t > 0, \tag{2}$$

[7] Such dependencies are, however, emphasised in section 2.2 where we discuss issues of identification.

where

$$A = \begin{pmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & I \\ A_0 & A_1 & A_2 & \dots & A_{p-2} & A_{p-1} \end{pmatrix}, \quad \phi(dt) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \zeta(dt) \end{pmatrix}.$$

The solution to (2) is given by

$$y(t) = e^{At}y(0) + \int_0^t e^{A(t-r)}\phi(dr), \quad t > 0, \tag{3}$$

where $y(0)$ denotes the vector of initial conditions and the matrix exponential is defined by its series expansion

$$e^{tA} = I + tA + \frac{1}{2!}(tA)^2 + \ldots = \sum_{j=0}^{\infty} \frac{(tA)^j}{j!}.$$

Noting that $y(th)$ contains the observable vector $x(th)$ the solution (3) can be manipulated to relate $y(th)$ to $y(th-h)$ and thereby $x(th)$ to $x(th-h)$. This is achieved by re-writing the system at the observation points as

$$\begin{aligned} y(th) &= e^{Ath}y(0) + \int_0^{th-h} e^{A(th-r)}\phi(dr) + \int_{th-h}^{th} e^{A(th-r)}\phi(dr) \\ &= e^{Ah}\left[e^{A(th-h)}y(0) + \int_0^{th-h} e^{A(th-h-r)}\phi(dr)\right] + \int_{th-h}^{th} e^{A(th-r)}\phi(dr). \end{aligned}$$

The term in square brackets is simply $y(th-h)$ which results in the following first-order stochastic difference equation for $y(th)$:

$$y(th) = Fy(th-h) + \epsilon_{th}, \quad t = 1, \ldots, T, \tag{4}$$

where $F = e^{Ah}$ and

$$\epsilon_{th} = \int_{th-h}^{th} e^{A(th-r)}\phi(dr)$$

is an i.i.d. random vector with mean vector zero and covariance matrix

$$\Sigma_\epsilon = \int_0^h e^{As}\Sigma_\phi e^{A's}ds,$$

$\Sigma_\phi dt$ being the covariance matrix of $\phi(dt)$.[8]

Although the system (4) implicitly embodies the dynamics of the observable vector $x_{th} = x(th)$ the remaining elements of $y(th)$ are unobservable. The Bergstrom approach derives the exact discrete time model by eliminating the unobservable elements from this system

[8] In fact, Σ_ϕ is an $np \times np$ matrix of zeros except for the $n \times n$ bottom right-hand corner block which is equal to Σ.

using appropriate substitutions.[9] This process results in the ARMA$(p, p-1)$ representation

$$x_{th} = F_1 x_{th-h} + \ldots + F_p x_{th-ph} + \eta_{th}, \quad t = p, \ldots, T, \tag{5}$$

where η_{th} is an MA$(p-1)$ process. Note that this equation holds only for period p onwards owing to the first available observation being $x_0 = x(0)$ (recall that we are assuming that x comprises purely stock variables at this point). It is, however, possible to derive an additional $p-1$ equations that relate $x_h, \ldots, x_{ph-h}$ to the lagged values and to x_0; see, for example, Theorem 2.2 of Bergstrom (1986) for the mixed sample case when $p = 2$, and Theorem 2 of Chambers (1999) also for the mixed sample case but for $p \geq 2$.

To see how this approach works in practice, consider the case where $p = 2$. The observable vector is $x(th)$ and the unobservable vector in this case is $Dx(th)$, the equations for which from (4) are

$$x(th) = F_{11} x(th-h) + F_{12} Dx(th-h) + \epsilon_{1,th}, \tag{6}$$

$$Dx(th) = F_{21} x(th-h) + F_{22} Dx(th-h) + \epsilon_{2,th}, \tag{7}$$

where the F_{ij} $(i, j = 1, 2)$ are the $n \times n$ submatrices of F and $\epsilon_{th} = (\epsilon'_{1,th}, \epsilon'_{2,th})'$. The objective is to eliminate $Dx(th-h)$ from (6) using the information in (7), and for this purpose Bergstrom (1983, Assumption 4) assumes that the matrix F_{12} is nonsingular. From (6) we obtain, using this assumption,

$$Dx(th-h) = F_{12}^{-1} \left[x(th) - F_{11} x(th-h) - \epsilon_{1,th} \right], \tag{8}$$

while lagging (7) by one period yields

$$Dx(th-h) = F_{21} x(th-h) + F_{22} Dx(th-2h) + \epsilon_{2,th-h}. \tag{9}$$

Substituting the right-hand-side of (8) for $Dx(th-h)$ in (9) and the one-period lag of (8) for $Dx(th-2h)$ in (9) results in

$$x_{th} = F_1 x_{th-h} + F_2 x_{th-2h} + \eta_{th}, \quad t = 2, \ldots, T, \tag{10}$$

where $F_1 = F_{11} + F_{12} F_{22} F_{12}^{-1}$, $F_2 = F_{12}[F_{21} - F_{22} F_{12}^{-1} F_{11}]$, and the disturbance vector is given by $\eta_{th} = \epsilon_{1,th} - F_{12} F_{22} F_{12}^{-1} \epsilon_{1,th-h} + F_{12} \epsilon_{2,th-h}$ which is clearly seen to be MA(1) due to ϵ_{th} being an i.i.d. process.

Although the ARMA(2,1) representation in (10) holds for $t = 2, \ldots, T$ it is possible to supplement it, for purposes of computing the unconditional likelihood function, with an equation that relates x_h to x_0. In the case of the second-order system considered here the relevant equation is given by (6) evaluated at $t = 1$, giving

$$x_h = F_{11} x_0 + F_{12} Dx(0) + \epsilon_{10}. \tag{11}$$

Note that this equation also includes the unobservable component $Dx(0)$, and there are two main ways of treating it. The first is to make an assumption about its value, an example being

[9] Wymer (1972) provided the first treatment of higher-order systems in the econometrics literature using the framework (1)–(4) but subsequently derived an approximate discrete model.

$Dx(0) = 0$, which implies that, at time $t = 0$, the system was in equilibrium. Alternatively the $n \times 1$ vector $Dx(0)$ can be treated as part of the unknown parameter vector whose value is estimated by maximisation of the likelihood function, although in this case it is not possible to obtain a consistent estimator of its value.[10]

As mentioned earlier, not all variables are observed as stocks, and so the above techniques have to be modified in the presence of flow variables or mixtures of stocks and flows. This is particularly important in macroeconometric modelling where many variables, such as consumers' expenditures and national income, are measured as flows. Early contributions dealing with the problems associated with flow variables can be found in Phillips (1974) and Wymer (1976). In subsequent work Bergstrom (1984, Theorem 8) presented an exact discrete time model for a first-order system while an exact discrete model for flow variables when $p = 2$ was derived by Bergstrom (1983, Theorem 3) and extended to the mixed sample case by Bergstrom (1986, Theorems 2.1 and 2.2).[11] In these cases the exact discrete time model can be shown to be an ARMA(p, p) system, the presence of flows increasing the order of the moving average disturbance by one. These results were subsequently extended to the general $p \geq 2$ case by Chambers (1999).

A feature of the results cited above is that all require an assumption of invertibility of certain matrices; for example, Bergstrom (1983) requires A_0 to be nonsingular in addition to F_{12}. The nonsingularity of A_0 rules out important cases such as unit roots and cointegration (see section 2.4), but can be relaxed as follows. Our demonstration applies to the case $p = 2$ but can be generalised to larger values of p. Recalling the definition of the observed flow variables, x^f_{th}, we can integrate (6) and (7) over the interval $(th - h, th]$ to obtain

$$x^f_{th} = F_{11}x^f_{th-h} + F_{12}z_{th-h} + v_{1,th}, \tag{12}$$

$$z_{th} = F_{21}x^f_{th-h} + F_{22}z_{th-h} + v_{2,th}, \tag{13}$$

where we have defined

$$z_{th} = \int_{th-h}^{th} Dx^f(r)dr = x^f(th) - x^f(th - h),$$

$$v_{th} = \begin{pmatrix} v_{1,th} \\ v_{2,th} \end{pmatrix} = \int_{th-h}^{th} \int_{s-h}^{s} e^{A(s-r)}\phi(dr)ds.$$

The vector z_{th} is unobservable and can be eliminated from the system using the same steps that led to (10), the result being

$$x^f_{th} = F_1 x^f_{th-h} + F_2 x^f_{th-2h} + \eta^f_{th}, \quad t = 2, \ldots, T, \tag{14}$$

where F_1 and F_2 are defined following (10) and $\eta^f_{th} = v_{1,th} - F_{12}F_{22}F_{12}^{-1}v_{1,th-h} + F_{12}v_{2,th-h}$ is now an MA(2) process which follows by noting that v_{th} can be written under the white noise assumption as the sum of a pair of single intervals with respect to $\zeta(dr)$ over the intervals $(th - 2h, th - h]$ and $(th - h, th]$; details can be found in McCrorie (2000). Although the autoregressive matrices remain the same functions of the underlying parameters as in the

[10] Note that this inconsistency arises owing to no new information on $Dx(0)$ becoming available as $T \to \infty$.

[11] Bergstrom (1986) also includes results for a system that contains exogenous stock and flow variables.

case of stock variables, the presence of flows affects the serial correlation properties of the disturbance vector, increasing the moving average order by one, a feature which needs to be incorporated in any estimation algorithm.

Although autoregressive models, in both discrete and continuous time, dominate the time series econometrics literature, there has been considerable interest in continuous time ARMA (CARMA) processes in the statistics literature, where the focus has been on state space approaches rather than exact discrete time representations. Results on maximum likelihood estimation based on an appropriate state space model are contained in Zadrozny (1988) while a survey of recent results on CARMA processes can be found in Brockwell (2014). It is, however, possible to derive an exact discrete time model corresponding to a CARMA system. Chambers and Thornton (2012) extend (1) to the CARMA(p,q) system

$$D^p x(t) = A_{p-1}D^{p-1}x(t) + \ldots + A_0 x(t) + u(t) + \Theta_1 Du(t) + \ldots + \Theta_q D^q u(t),\ t > 0, \quad (15)$$

where $u(t)$ is an $n \times 1$ continuous time white noise process and $A_0, \ldots, A_{p-1}$ and $\Theta_1, \ldots, \Theta_q$ are $n \times n$ matrices of coefficients.[12] The interpretation of a white noise process in continuous time can be problematic (see, for example, the discussion and results in Bergstrom, 1984) but the interpretation of $u(t)$ in (15) is that it satisfies $E[u(t)] = 0$ and, for $t_2 > t_1$, has autocovariance properties

$$E\left[\int_{t_1}^{t_2} u(r)dr \int_{t_1}^{t_2} u(s)'ds\right] = \Sigma\left(t_2 - t_1\right),$$

$$E\left[\int_{t_1}^{t_2} u(r)dr \int_{t_1}^{t_2} u(\tau + s)'ds\right] = 0, \quad |\tau| > t_2 - t_1,$$

where Σ is an $n \times n$ positive definite symmetric matrix.

The presence of the MA component in (15) means that a different state space form is more useful in deriving the exact discrete model than the one defined in (2). Chambers and Thornton (2012) employed the state space representation used by Zadrozny (1988) in which the $np \times 1$ state vector is defined as $w(t) = [w_1(t)', \ldots, w_p(t)']'$ and with $w_1(t) = x(t)$. The state space form is based on the following set of p equations in the derivatives of the components of $w(t)$, given by

$$\begin{aligned}
Dw_1(t) &= A_{p-1}w_1(t) + w_2(t) + \Theta_{p-1}u(t), & (16)\\
Dw_2(t) &= A_{p-2}w_1(t) + w_3(t) + \Theta_{p-2}u(t), & (17)\\
&\vdots \\
Dw_{p-1}(t) &= A_1 w_1(t) + w_p(t) + \Theta_1 u(t), & (18)\\
Dw_p(t) &= A_0 w_1(t) + u(t), & (19)
\end{aligned}$$

in which we define $\Theta_j = 0$ for $j > q$. Combining the expressions for $Dw_1(t), \ldots, Dw_p(t)$ above, the relevant state space form can be written

$$Dw(t) = Cw(t) + \Theta u(t), \quad (20)$$

[12] The coefficient matrix multiplying $u(t)$ is set to an identity in order to identify the parameters of the model in view of $u(t)$ having covariance matrix Σ.

where

$$C = \begin{bmatrix} A_{p-1} & I & 0 & \dots & 0 \\ A_{p-2} & 0 & I & \dots & 0 \\ \vdots & & & & \vdots \\ A_1 & 0 & 0 & \dots & I \\ A_0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad \Theta = \begin{bmatrix} \Theta_{p-1} \\ \Theta_{p-2} \\ \vdots \\ \Theta_1 \\ I \end{bmatrix}.$$

Utilising this state space model Chambers and Thornton (2012) show that the exact discrete time model for a vector of stock variables is of ARMA$(p, p-1)$ form while for a vector of flow variables or mixed sample data it is of ARMA(p, p) form. The presence of the continuous time MA disturbance therefore does not affect the MA order of the exact discrete time model. This means, in effect, that there are additional parameters in the CARMA model that can be used to pick up the dynamics in the discrete time model that are not present in a CAR representation, a feature that has been shown to have empirical content by Chambers and Thornton (2012).

More recently, Thornton and Chambers (2017) have shown that exact discrete time representations corresponding to CARMA systems are not unique.[13] The discrete time representations for CAR(p) systems with mixed sample data, developed in Bergstrom (1983) and in Chambers (1999), rely on differencing the stock variables and are of ARMA(p, p) form. Once the stock variables are re-integrated (or 'un-differenced'), these representations correspond to a discrete time ARMA$(p + 1, p)$ process. Thornton and Chambers (2017), however, work with an augmented state space form[14] that more naturally incorporates both stock and flow variables and show that the differencing of the stock variables identifies the representation among a wider class of ARMA$(p + 1, p)$ processes and that the more parsimonious ARMA(p, p) is also among this class.

2.2. Identification

To a large extent, we motivated the formulation of continuous time models as linear stochastic differential systems because equispaced data generated by such systems satisfy ARMA specifications that are typical in time series analysis but whose parameters, unlike those in naïvely-specified discrete-time models, are not tied to the sampling interval. The principal counterpoint to this advantage of estimating the parameters of structural continuous time models on the basis of discrete data is that one can 'join up the dots,' as Robinson (1992) described it, in an uncountably infinite number of ways. The problem is multivariate in character and can be illustrated using the following simple example for a stock variable.[15] Suppose that the $n \times 1$ finite-variance vector $x(t)$ satisfies the stochastic differential equation

[13] Hence the presence of the phrase '*an* exact discrete time representation' rather than '*the* exact discrete time representation' in the title of this chapter.

[14] The state space form in (20) is augmented by an additional n^f elements in a vector $y_0(t)$ that corresponds to the aggregated or observed flow variables.

[15] This identification problem is therefore different in nature and on top of the classical identification problem which seeks to avoid observational equivalence through model and estimator choice; see, for example, Chambers and McCrorie (2006). In open systems, namely systems involving exogenous variables, the solution of the stochastic differential equation depends on a continuous time record of the exogenous variables and so some sort of approximation of the time paths is necessary to achieve identification; see, in particular, Bergstrom (1986), Hamerle, Nagl and Singer (1991), Hamerle, Singer and Nagl (1993) and McCrorie (2001) for explicit discussion of this issue.

system

$$dx(t) = A(\theta)x(t)dt + \zeta(dt), \quad t > 0, \tag{21}$$

subject to the initial condition $x(0) = y_0$, where A is an $n \times n$ matrix whose elements are now explicitly assumed to be known functions of a $p \times 1$ vector θ of unknown parameters ($p \leq n^2$), y_0 is a non-random $n \times 1$ vector, and $\zeta(dt)$ is an uncorrelated vector random measure of the type described in section 2.1 with covariance matrix $\Sigma(\mu)dt$, the elements of Σ being known functions of a $q \times 1$ vector μ of unknown parameters ($q \leq n(n+1)/2$). The exact discrete time model is obtained from the solution of (21) subject to the initial condition, giving a sequence of equispaced discrete time data $x(0), x(h), \ldots, x(Th)$ that satisfies the stochastic difference equation system

$$x(th) = F(\theta)x(th - h) + \epsilon_{th}, \quad t = 1, \ldots, T, \tag{22}$$

where $F(\theta) = e^{A(\theta)h}$ and ϵ_{th} is white noise with covariance matrix

$$\Omega_\epsilon(\theta, \mu) = E(\epsilon_{th}\epsilon'_{th}) = \int_0^h e^{A(\theta)r}\Sigma(\mu)e^{A(\theta)'r}dr;$$

see Bergstrom (1984, Theorem 3).

In the context of (21), the identification problem relates directly to the fact that there are, in principle, many different matrices that share the same exponential F in (22); see, for example, Phillips (1973), Hansen and Sargent (1983) and Hamerle, Singer and Nagl (1993). These matrices are *aliases* of A in the sense that, through taking the place of A in (21), they generate the same equidistant discrete time data. The aliasing problem of identifying structural continuous time parameters on the basis of discrete time data is clearly more severe than simply identifying the parameters of discrete time models (but, to reiterate, there is a trade-off in that naïvely-specified discrete-time models suffer from a lack of time invariance). If Gaussianity is assumed, the problem in the context of (21) is to find a necessary and sufficient condition such that the pair $[A(\theta), \Sigma(\mu)]$ is identifiable in $[F(\theta), \Omega_\epsilon(\theta, \mu)]$. In any particular application, the forms of A and Σ are heavily governed by the role of the parameter vectors θ and μ, although for the purpose of simplifying the discussion that follows, the dependence of A and Σ on θ and μ will be suppressed.

McCrorie (2003) offered a framework for the identification problem by considering the following Hamiltonian matrix M that allows the pair $[A, \Sigma]$ to be treated together: if

$$M = \begin{pmatrix} -A & \Sigma \\ 0 & A' \end{pmatrix},$$

then, as an application of Van Loan (1978, Theorem 1),

$$e^{Mh} = \begin{pmatrix} F^{-1} & F^{-1}\Omega_\epsilon \\ 0 & F' \end{pmatrix}$$

The following theorem, which is a consequence of Theorem 2 of Culver (1966), contains the basic result on identification in terms of when the matrix exponential mapping is bijective in general.

Theorem (McCrorie, 2003). For the prototypical model (21), $[A, \Sigma]$ is identifiable in $[F, \Omega_\epsilon]$ if the eigenvalues of M are strictly real and no Jordan block of M belonging to any eigenvalue appears more than once.

Note that the eigenvalues of M are simply the eigenvalues and reverse eigenvalues of A, and so if A has no complex eigenvalues and there is no confluence in its eigenvalues, the aliasing problem reduces essentially to a univariate problem involving the exponential function which, when viewed as real-valued, is bijective. Unfortunately, both restrictions are not generally appropriate for economic time series: they rule out plausible cyclical behaviour resulting from complex eigenvalues and plausible trend behavior resulting from multiple unit roots (multiple zero eigenvalues of A). In the complex eigenvalue case, several authors achieve identification through additional restrictions: Phillips (1973) uses Cowles Commission type restrictions (see also Blevins, 2017) and Hansen and Sargent (1983) show there are restrictions inherent in the requirement that Ω_ϵ be positive semidefinite. Hansen and Sargent (1991) use cross-equation restrictions implied by the rational expectations hypothesis. Bergstrom, Nowman and Wymer (1992) and Bergstrom and Nowman (2007) use prior bounds on the parameters as a means of achieving identification, in the way researchers do for large-scale structural VAR models today. The results of Hansen and Sargent (1983) show that without importing *a priori* restrictions beyond the problem in hand, identification can only be local; see Appendix 1 of McCrorie (2009) for some examples. In practice, one has *jointly* to solve the aliasing identification problem and the classical identification problem of avoiding observational equivalence through model and estimator choice. In the context given here, the general problem relates not to the matrices A and Σ but to the underlying parameter vectors θ and μ. The incorporation of exogenous variables in open systems can be useful (e.g. Hamerle, Singer and Nagl, 1993; Bergstrom, Nowman and Wymer, 1992), as can the aspect that the matrices in terms of the underlying parameter vectors are often heavily restricted. Nevertheless, finding necessary and sufficient conditions to solve the identification problem for estimating continuous time models on the basis of discrete data remains open even for the most basic of models.

2.3. Granger causality

Formulating a structural model in continuous time offers a means of resolving the problem that discrete time models, whose estimated parameters are tied to the sampling frequency, do not readily lend themselves to economic interpretation. A parallel problem that has also been downplayed in the econometrics literature is the tendency for naïvely-specified discrete time models to generate spurious Granger causality relationships when the time intervals in which the data are generated are finer than the sampling interval.[16] To define (global) Granger non-causality between two variables $x_1(t)$ and $x_2(t)$, let $I_j(t)$ $(j = 1, 2)$ denote the sigma algebra generated by $x_j(t)$ up to time t (this is interpreted as an information set), let $\bar{I}(t)$ denote all other information up to time t, and let $E(A|B)$ denote the conditional expectation of A given B. Then x_2 *does not Granger cause* x_1 if

$$E(x_1(t+k)|I_1(t), I_2(t), \bar{I}(t)) = E(x_1(t+k)|I_1(t), \bar{I}(t)) \quad \text{for all } t \text{ and } k > 0; \qquad (23)$$

see Florens and Fougère (1996) and Comte and Renault (1996). If the above condition does

[16] McCrorie and Chambers (2006, Section 3.1) outline and discuss the concept of Granger causality in the context of continuous and discrete time models.

not hold then x_2 is said to Granger cause x_1. A brief survey and discussion of the literature of Granger causality in continuous time can be found in McCrorie and Chambers (2006, Section 3).

In the context of temporal aggregation a coarsely sampled process, omitting information useful for predicting an economic time series, will exhibit bidirectional Granger causality with another coarsely sampled process provided that they are correlated, even if there is only unidirectional causality in the finer time interval. Inferences made about the underlying behaviour of economic agents from observed time series can, therefore, be distorted. For example, Christiano and Eichenbaum (1987) find evidence for the money stock Granger-causing output with quarterly U.S. data that seems to be overturned when moving to a finer sampling interval. Some authors, for example Marcellino (1999) and Breitung and Swanson (2002), have tried to approach the temporal aggregation problem through the lens of fixed-interval time aggregation; however, this approach relies on constructing corrections to estimates through knowing the time unit in which the data are generated. Otherwise, a distortional effect owing to temporal aggregation will remain.[17]

Specifying a structural continuous time model allows *a priori* restrictions to be imposed on the observed discrete data independently of the sampling interval, enabling Granger causality relationships to be preserved, and thereby facilitates obtaining efficient estimates of the structural parameters that are devoid of temporal aggregation bias. Such considerations matter materially in empirical work. For example, Harvey and Stock (1989) find evidence, using U.S. data, of the money stock not Granger-causing output on the basis of a continuous time model but obtain a strong reversal of this conclusion when temporal aggregation is ignored in discrete-time VARs. McCrorie and Chambers (2006) also consider the issue of money-income causality in discrete time models where the temporal aggregation restrictions were imposed exactly, approximately and not at all. They find that accounting for temporal aggregation restrictions can have an important bearing on Granger causality tests, even when the restrictions are only approximately imposed. In an application to exchange rates, Renault, Sekkat and Szafarz (1998) used a continuous time model and the methods of this chapter to distinguish between 'true' and 'spurious' causality, and on the basis of their data suggested that there was a 'discrete-time illusion' of spurious causality observed between the German mark and the Swiss franc at certain sampling frequencies.

The above discussion motivates formulating continuous time models as a way of countering the problem that some observed Granger-causality relationships in naïvely-specified discrete time models are spurious. In practice, there exists a *trade-off* between preserving *a priori* information on Granger causality relationships in estimation with solving the problem of identifying the parameters of a structural continuous time model on the basis of discrete time data as discussed in section 2.2. Both issues are in the background regardless of the model formulated. For example, naïvely specifying a discrete time model on its own, common throughout econometric time series analysis, is insufficient as it gives no reference point to assess whether the magnitude of temporal aggregation is important.

2.4. Nonstationarity

Economic time series data are inherently nonstationary, and the nonstationarity can

[17] Thornton and Chambers (2013) provide a recent discussion of temporal aggregation in macroeconomics with continuous time models in view.

manifest itself in a variety of forms. A second-order stationary time series is one for which the mean, variance and autocovariances are time-independent. Examination of the solution to the state space representation of the continuous time system given in (3) shows immediately that the mean depends on time because $E[y(t)] = e^{At}y(0)$, assuming $y(0)$ is fixed. This is the form of nonstationarity referred to in the title of Bergstrom (1985). However, in recent years nonstationarity has come to be associated with a different concept, namely that of unit roots and stochastic trends, which are consistent with the earlier observation of Granger (1966) concerning the shape of the spectral density function at the origin.

A discrete time process, x_{th} $(t = 1, \ldots, T)$, is said to have a unit root if it has the representation

$$\Delta_h x_{th} = u_{th}, \quad t = 1, \ldots, T, \tag{24}$$

where $\Delta_h x_{th} = x_{th} - x_{th-h}$ and u_{th} is a second-order stationary random process. Solving the difference equation from an initial value x_0 yields the representation for the level process in the form

$$x_{th} = x_0 + \sum_{j=1}^{t} u_{jh}, \quad t = 1, \ldots, T, \tag{25}$$

where the partial sum of the stationary process u_{th} represents the stochastic trend. One way of thinking about a unit root is that the process requires differencing once to become stationary, as in (24), while the stochastic trend representation (25) leads to the levels process being described as integrated of order one, often denoted I(1).

In continuous time the equivalent representation to (24) is

$$Dx(t) = u(t), \quad t > 0, \tag{26}$$

where $u(t)$ is a second-order stationary continuous time process and $x(0)$ will be taken to be fixed. In this case the process $x(t)$ requires differentiating once to become stationary and the stochastic trend representation for the level is given by

$$x(t) = x(0) + \int_0^t u(r)dr, \quad t > 0, \tag{27}$$

assuming the integral (which represents the continuous time stochastic trend) exists. If $x(t)$ is observed as a discrete time process at integer values of t at intervals of length h then integrating (26) once over the interval $(th - h, th]$ reveals that

$$x(th) = x(th - h) + \int_{th-h}^{th} u(r)dr$$

and hence the discrete time process has a unit root. This is also true of an observed flow variable obtained by a further integration of the model above.

Following Phillips (1987) a large literature has emerged on unit root processes in discrete time and much effort has been expended in the search for tests for a unit root that have good properties. Many economic time series have been found to display unit root-type properties but one of the challenges facing economics in the mid-1980s was how to reconcile economic theory with these apparent features. In particular, individual series with unit roots can wander freely over time, driven by the stochastic trends, whereas much of economics implies the existence of stable relationships among variables (an example being consumers' expenditure and income). The solution to this apparent dichotomy, proposed by Engle and Granger (1987), was the concept of cointegration. An $n \times 1$ vector, x_{th}, of I(1) series is said to be cointegrated if there exist a set of $1 \leq r < n$ linear combinations of the form $\beta' x_{th}$ that are stationary, where β is an $n \times r$ matrix of cointegrating parameters whose columns are the r cointegrating vectors. Cointegration has subsequently become an essential concept in the analysis of multivariate nonstationary economic time series.

In terms of continuous time processes, Phillips (1991) showed that a vector process that is cointegrated in continuous time is also cointegrated in terms of the discrete time observations.[18] This is an important result because it implies that discrete time methods can be used to test for cointegration even if the researcher is interested in formulating a model in continuous time. If evidence of r cointegrating vectors is found, let $m = n - r$ and partition $x(t) = [x_1(t)', x_2(t)']'$, where $x_1(t)$ is $r \times 1$ and $x_2 t)$ is $m \times 1$. Then there exists an $r \times m$ matrix, B, of cointegrating vector such that $x_1(t) - Bx_2(t)$ is a stationary continuous time process. Note that these cointegrating relationships have been normalised on $x_1(t)$, which is an identification condition. The continuous time model can then be represented in terms of an error correction model (ECM) of the form

$$Dx(t) = -JAx(t) + u(t), \quad t > 0, \tag{28}$$

where $J = [I_r, 0_{r\times m}]'$, $A = [I_r, -B]$ and $u(t)$ is a stationary process. The ECM representation (28) embodies two key features of the cointegrated system. The first r equations are of the form

$$Dx_1(t) = -[x_1(t) - Bx_2(t)] + u_1(t), \quad t > 0,$$

in which x_1 is responding to the disequilibrium (or error) depicted by $x_1(t) - Bx_2(t)$. Such systems are often motivated by $x_1(t) = Bx_2(t)$ representing an equilibrium or optimal level of x_1 given the level of x_2. The remaining m equations in (28) are the stochastic trends driving the system; they are given by

$$Dx_2(t) = u_2(t), \quad t > 0,$$

subject to an initial value $x_2(0)$.

In continuous time cointegrated systems of the form (28) interest centres on estimation of the matrix B. Equispaced discrete time observations generated by this system satisfy

$$x(th) = e^{-JAh}x(th-h) + v(th), \quad v(th) = \int_{th-h}^{th} e^{-JA(th-r)}u(r)dr, \quad t = 1, \ldots, T.$$

Using the fact that $AJ = I_r$ the infinite series representation for the matrix exponential can be used to show that $e^{-JAh} = I_n - fJA$ where $f = 1 - e^{-h}$. It then follows that $x(th)$ satisfies the discrete time ECM

$$\Delta_h x(th) = -JAx(th-h) + w(th), \quad w(th) = v(th) + e^{-h}JAx(th-h), \quad t = 1, \ldots, T,$$

[18] Stock (1987) had earlier provided an example that cointegration as a property was invariant to temporal aggregation.

where $w(th)$ is a stationary disturbance vector in view of $Ax(th) = x_1(th) - Bx_2(th)$ being stationary. Phillips (1991) recommended the use of spectral regression estimators that treat the dynamics nonparametrically. Such methods exploit the stationary nature of $u(t)$ to the full without requiring any particular parametric form for the dynamics, and were shown to have good finite sample properties in the simulation study of Chambers (2001). Frequency domain methods can also be used in cointegrated systems in which the dynamics are modelled parametrically, for example in CAR(p) models such as (1) that embed cointegration by setting $A_0 = CA$, where A is defined following (28) and C is an $n \times r$ matrix of rank r. Chambers and McCrorie (2007) show that maximisation of a frequency domain likelihood function leads to estimates of B that are asymptotically mixed normal and to estimates of the autoregressive parameters that govern the dynamics that are asymptotically normal. As is the case with cointegrated systems in discrete time the estimates of B converge to the limiting distribution at rate T while those of the autoregressive parameters converge at rate $\sqrt{T}$. The exact discrete model corresponding to a first-order cointegrated system with mixed sample data was derived by Chambers (2009) and such models can be estimated based on the time domain Gaussian likelihood outlined earlier.[19] The effects of sampling frequency in the context of cointegrated continuous time CAR systems were also analysed by Chambers (2011).

2.5. Mixed frequency data

Time series data in economics are available at a variety of frequencies. Observations on macroeconomic aggregates, such as consumers' expenditure, investment, and national income, are typically available quarterly; variables such as the money supply and price indices used to compute measures of inflation are usually observed monthly; while financial variables, such as asset prices (interest rates, exchange rates, stock prices etc.) can be observed almost continuously but daily closing prices are often used. The extant approach to dealing with observations at different frequencies is to aggregate all variables to the lowest frequency, thereby potentially throwing away information contained in the high frequency observations that could be exploited for gains in modelling. For example, it might be possible to use high frequency financial variables to predict fluctuations in real economic activity before the low frequency observations are available. In recent years a number of advances in the analysis of mixed frequency data have been made and the topic has assumed added significance following the financial crisis of 2008.

In the context of continuous time models an often overlooked but nevertheless important contribution that incorporates observations at different frequencies was made by Zadrozny (1988). He considered the general problem of estimating a CARMA(p, q) system with mixed sample data available at mixed frequencies and recommended the use of state space forms and the Kalman filter for constructing the Gaussian likelihood function. More recently, and in keeping with the exact discrete time modelling approach, Chambers (2016) derived the exact discrete model corresponding to a CAR(1) system with mixed sample data observed at mixed frequencies. Suppose, for simplicity, that there are two vectors of stock variables, a low frequency one, x_2 ($n_2 \times 1$), observed at unit intervals of time, and a high frequency vector, x_1 ($n_1 \times 1$), observed at time intervals of length $0 < h < 1$ where it is convenient to

[19] Other time domain approaches to cointegrated models in continuous time can be found in Comte (1999) and Corradi (1997).

assume that $k = h^{-1}$ is an integer. For example, if x_2 is observed quarterly and x_1 monthly then $h = 1/3$ and x_1 is observed $k = 3$ times more frequently than x_2. Then, for each integer t, the $(kn_1 + n_2) \times 1$ vector

$$X_t = [x'_{1t}, x'_{1,t-h}, x'_{1,t-2h}, \ldots, x'_{1,t-(k-1)h}, x'_{2t}]', \quad t = 1, \ldots, T,$$

can be defined. The underlying continuous time model is assumed to be a CAR(1) system in the $n \times 1$ vector $x(t) = [x_1(t)', x_2(t)']'$ of the form

$$dx(t) = Ax(t)dt + \zeta(dt), \quad t > 0,$$

where $\zeta(dt)$ is defined following (1). The objective is to use the mixed frequency data to estimate the $n \times n$ matrix A and the $n(n+1)/2$ elements of the covariance matrix, Σ, of $\zeta(dt)$. Theorem 1 of Chambers (2016)[20] shows that the discrete time observations satisfy the exact discrete time model

$$\begin{aligned}
x_{1t} &= B_{11,1}x_{1,t-h} + \ldots + B_{11,k}x_{1,t-1} + B_{12,0}x_{2,t-1} + \eta_{1t}, \\
x_{1,t-h} &= B_{11,1}x_{1,t-2h} + \ldots + B_{11,k-1}x_{1,t-1} + B_{12,1}x_{2,t-1} + \eta_{1,t-h}, \\
&\vdots \qquad\qquad\qquad \vdots \\
x_{1,t-(k-1)h} &= B_{11,1}x_{1,t-1} + B_{12,k-1}x_{2,t-1} + \eta_{1,t-(k-1)h}, \\
x_{2t} &= \sum_{j=1}^{k} B_{21,j}x_{1,t-jh} + B_{22}x_{2,t-1} + \eta_{2t},
\end{aligned}$$

where the $B_{ij,k}$ matrices are of dimension $n_i \times n_j$ $(i, j = 1, 2)$ and the $(kn_1 + n_2) \times 1$ vector

$$\eta_t = [\eta'_{1t}, \eta'_{1,t-h}, \eta'_{1,t-2h}, \ldots, \eta'_{1,t-(k-1)h}, \eta'_{2t}]'$$

is a vector white noise process. It is important to stress that all of the autoregressive matrices in the mixed frequency discrete time representation are only functions of the elements of the matrix A while the covariance matrix of η_t depends only on A and Σ. By way of comparison a discrete time vector autoregression in the vector X_t would be significantly over-parameterised.

Similar exact discrete time models can be derived for the cases where both the high and low frequency observations are on flow variables and where they are mixtures of stocks and flows. The main difference when flow variables are present is that the disturbance vector becomes an MA(1) process but the parsimony over unrestricted VAR and VARMA systems remains. Simulation results in Chambers (2016) for stationary and cointegrated systems show that utilising the mixed frequency data reduces bias and mean squared error of Gaussian estimates compared with the situation where high frequency variables are aggregated to the low frequency. Furthermore, in an empirical application testing long run purchasing power parity restrictions between the UK and the US, inferences are found to be unfavourable to the restrictions when using the information in daily frequency exchange rates but the restrictions are not rejected when the exchange rates are aggregated to weekly and monthly frequencies. A possible explanation for this finding is that the estimates of the two key

[20] The model considered by Chambers (2016) also includes a vector of intercepts and deterministic trends.

parameters of interest have large standard errors using the aggregated series but are more precisely determined when using the high frequency data.

2.6. Gaussian estimation using an exact discrete model

The exact discrete time model in the form of (10), allied with an additional set of $p-1$ conditions relating the initial observations to the initial state vector in the case of stock variables or p such conditions in the case of flows or a mixed sample, provides a basis for the construction of the likelihood function. It is usually assumed that the $nT \times 1$ vector $\eta = (\eta_h', \ldots, \eta_{Th}')'$ is Gaussian with mean vector zero and covariance matrix $\Omega_\eta = E(\eta\eta')$, which is equivalent to specifying $\zeta(dt)$ in (1) to be the increment in a Brownian motion process. Under such an assumption the log-likelihood is of the form

$$\log L(\theta) = -\frac{nT}{2}\log 2\pi - \frac{1}{2}\log|\Omega_\eta| - \frac{1}{2}\eta'\Omega_\eta^{-1}\eta, \tag{29}$$

where θ denotes the parameter vector of interest (i.e. the elements of the autoregressive matrices $A_0, \ldots, A_{p-1}$ and the covariance matrix Σ). Section 3 discusses computational aspects associated with (29).

The asymptotic properties of estimates obtained by maximising (29) depend, of course, on the set of assumptions made concerning the model (1). Bergstrom (1983) provided a set of conditions that ensures that the vector, $\hat{\theta}$, that maximises (29) is almost surely consistent and, furthermore, that $\sqrt{T}(\hat{\theta}-\theta)$ is asymptotically normal and efficient in the Cramer sense. These conditions include such things as identification of θ in a closed bounded set Θ over which maximisation takes place; stationarity and ergodicity of $x(t)$; and continuity and differentiability of the autoregressive matrices and covariance matrix of (1) in cases where the elements of these matrices may depend, possibly nonlinearly, on an underlying parameter vector of smaller dimension. The issue of identification has an added dimension in continuous time models owing to the phenomenon of aliasing which was discussed in section 2.2.

In finite samples the problem of estimation bias has the potential to beset *all* Gaussian/maximum likelihood (ML) methods including those based on the exact discrete time model. It is particularly relevant when estimating mean reversion parameters, as demonstrated in Phillips and Yu (2005) and Yu (2012). In a sampling experiment using a common interest rate model, Phillips and Yu (2009) showed that the estimation bias can be more important than the bias arising from using an approximate rather than an exact solution of the continuous time model. Wang, Phillips and Yu (2011) decompose the overall bias into separate terms arising from estimation and from discretisation, finding that when using Euler and trapezoidal approximations to the exact discrete model, both approximate methods dominate the exact method for empirically realistic cases. They also show that the sign of the discretisation bias is opposite to that of the estimation bias in such cases, meaning that the bias in the approximate methods is less than for estimation based on the exact discrete model. In addition the asymptotic variance of the estimator based on the Euler approximation is smaller than for the ML estimator of the mean reversion parameter in the exact discrete model, supporting a conclusion that the Euler approximation would be preferred to ML estimation of the exact discrete model in certain circumstances, such as when mean reversion in a univariate linear diffusion is slow.

It should be borne in mind that the exact discrete model is the only model that exactly incorporates restrictions implied by economic theory and other *a priori* information on the observed discrete data, and methods have been proposed to reduce finite sample estimation bias. Phillips and Yu (2005, 2009) propose jackknife techniques and a simulation-based indirect inference method and show they are successful in reducing finite sample bias in univariate diffusion models. Jackknife methods can also be expected to work successfully in higher-order and multivariate continuous time models, as indicated by the results of Chambers (2013) for stationary autoregressions and Chambers and Kyriacou (2013) in unit root models. The application of such techniques to more general continuous time systems is worthy of further investigation.

2.7. Alternative approaches

Although we have emphasised the exact discrete time modelling approach to the estimation of continuous time systems it is not the only suitable method. As mentioned above Zadrozny (1988) has shown how Kalman filtering techniques can be used to compute the Gaussian likelihood function in CARMA systems with mixed sample and mixed frequency data that can also include exogenous variables. State space forms and the Kalman filter were also used in a sequence of contributions by Harvey and Stock (1985, 1988, 1989) that built on earlier work by Jones (1981) and focused on CAR systems that may contain integrated and cointegrated variables. Singer (1995) also proposed a filtering method and used analytic derivatives to facilitate computing the likelihood function. The evaluation of the likelihood function using the exact discrete model approach treats the entire observation vector simultaneously whereas the Kalman filter is a recursive method that is usually defined stepwise from observation to observation. However, both methods should produce the same value for the likelihood function. Bergstrom (1985) offers some comparison between the methods, as do Singer (2007) and Oud and Singer (2008) though for methods extended to deal with panel data.

The main advantage of these approaches is that it is not necessary to derive the full exact discrete time model, merely the first-order difference equation satisfied by the state vector that includes unobservable components as well as the observed variables. Another advantage of this approach is that the Kalman filter produces optimal estimates of the unobservable components of the state vector which may be of interest in certain applications. A disadvantage is that it is less readily comparable to alternative discrete time models, a shortcoming that is clearly not shared by the exact discrete time representation. Furthermore, Bergstrom (1985) provided some arguments as to why the exact discrete time approach has computational advantages over the Kalman filter approach, although no formal testing of these claims appears to have been conducted and will, no doubt, depend on a whole variety of factors.[21]

Alternative frequency domain methods can also be used to estimate stationary CARMA systems. The spectral density matrix of the continuous time process $x(t)$ in (15) is given by

$$F(\lambda) = \frac{1}{2\pi} A(-i\lambda)^{-1} \Theta(-i\lambda) \Sigma \Theta(i\lambda)' [A(i\lambda)']^{-1}, \quad -\infty < \lambda < \infty, \tag{30}$$

[21]Such factors include, but are not restricted to: the order of the continuous time system; the dimension of the vector $x(t)$; the sample size; the way in which the likelihoods are programmed; and, the optimisation algorithm used.

where $i^2 = -1$,

$$\begin{aligned} A(z) &= z^p I_n - A_{p-1} z^{p-1} - \ldots - A_1 z - A_0, \\ \Theta(z) &= I_n + \Theta_1 z + \ldots + \Theta_{q-1} z^{q-1} + \Theta_q z^q. \end{aligned}$$

Assuming that $x(t)$ is comprised of stock variables the spectral density matrix of the discretely observed vector x_{th} is given by

$$F^d(\lambda) = \frac{1}{h} \sum_{j=-\infty}^{\infty} F\left(\frac{\lambda + 2\pi j}{h}\right), \quad -\pi < \lambda \leq \pi,$$

the so-called folding formula. Robinson (1993) provides formulae that enable $F^d(\lambda)$ to be computed exactly so that a frequency domain version of the Gaussian likelihood function (or Whittle likelihood) can be constructed. Flow variables are also easily handled within this framework, as are mixed samples. Suppose $x(t) = [x^s(t)', x^f(t)']'$ and we partition $F(\lambda)$ as

$$F(\lambda) = \begin{pmatrix} F^{ss}(\lambda) & F^{sf}(\lambda) \\ F^{fs}(\lambda) & F^{ff}(\lambda) \end{pmatrix}.$$

Then the spectral density matrix of the continuous time process

$$X(t) = \begin{pmatrix} x^s(t) \\ \frac{1}{h}\int_{t-h}^{t} x^f(r) dr \end{pmatrix}$$

is given by Robinson (1993) as

$$F_X(\lambda) = \begin{pmatrix} F^{ss}(\lambda) & \frac{1 - e^{-ih\lambda}}{ih\lambda} F^{sf}(\lambda) \\ \frac{e^{ih\lambda} - 1}{ih\lambda} F^{fs}(\lambda) & \frac{4\sin^2 h\lambda/2}{h^2\lambda^2} F^{ff}(\lambda) \end{pmatrix}, \quad -\infty < \lambda < \infty.$$

The terms multiplying components of the spectral density matrix involving flow variables arise through the frequency response function of the integral determining the observed process (and the squared frequency response function for $F^{ff}(\lambda)$). The spectral density of the process observed at discrete points in time i.e. for $X(th)$ $(t = 1, 2, \ldots)$ is then subject to the folding formula yielding

$$F_X^d(\lambda) = \frac{1}{h} \sum_{j=-\infty}^{\infty} F_X\left(\frac{\lambda + 2\pi j}{h}\right), \quad -\pi < \lambda \leq \pi.$$

Fourier methods for the estimation of even more general continuous time systems were earlier proposed by Robinson (1976).

3. Computational issues

Except in the simplest cases, estimates of the parameters of continuous time models do not have closed form solutions and typically require optimisation using programmable statistical software such as R, Matlab or Gauss. Fortunately, the growth in computing power has expanded the scope and the dimension of feasible models, provided the sparse nature of many of the matrices involved in computing the likelihood and the possibility of in-sample convergence is exploited.

Firstly, the translation of an autoregressive model in continuous time to a discrete time model free of dependence on any sampling frequency involves the calculation of a matrix exponential, as in equation (3), and functions thereof. Owing to results by Van Loan (1978), the functions of the exponential can be computed as products of submatrices of a single, larger dimensional matrix exponential. Chambers (1999), McCrorie (2000) and Thornton and Chambers (2016) provide expressions pertaining to the exact discrete time model, while Harvey and Stock (1985) and Zadrozny (1988) provide similar expressions for application of the Kalman filter.

Moler and Van Loan (1978), in a celebrated article in the numerical analysis literature,[22] showed that computation of the matrix exponential is a notoriously ill-conditioned problem, to the extent that of nineteen methods considered, only three or four were potentially suitable in general, including a scaling and squaring method that employs Padé approximation to the scalar exponential (see Higham, 2009). Jewitt and McCrorie (2005) discuss the computational issues behind computing matrix exponentials and their functions with continuous time econometrics in view. Standard methods are not always robust. For example, taking the partial sums of the Taylor series following equation (3) can be ill-conditioned because round-off error can propagate in computing higher and higher powers in a way that eventually dominates analytical convergence. A popular alternative is to exploit an eigenvalue decomposition when A is diagonalisable i.e. when A is similar to a diagonal matrix Λ containing the eigenvalues of A. If $A = Q\Lambda Q^{-1}$ then $e^A = Qe^{\Lambda}Q^{-1}$, where e^{Λ} is, conveniently, a diagonal matrix whose elements are exponentials of the corresponding elements of Λ. The method relies, however, on an *a priori* assumption that the matrix A is diagonalisable, which is inconsistent with the property of cointegration that economic data plausibly satisfy. It is also possible that Q itself is ill-conditioned; see, for example, Higham and Al-Mohy (2010, Section 4). The main recommendation of Jewitt and McCrorie (2005) is that, for the type of matrices liable to be seen in econometric modelling, the problem is not likely to be ill-conditioned should any of the three standard methods discussed therein, including the scaling and squaring method also recommended by Zadrozny (1988), be used and supported by calculations made to at least standard IEEE double precision.

Hereafter, computation of the likelihood using the exact discrete representation diverges from calculation using the Kalman filter. The Kalman filter may be applied to equation (4) in association with an observation equation that synthesises the observed series, x_{th}, from the state vector, $y(th)$. Well known methods have been developed to cope with irregularly spaced date and with observation noise; see, for example, Harvey (1989). The likelihood is often evaluated using the T prediction error vectors for the observed series, which, being optimal linear predictions, are uncorrelated. Such routines have the advantage/incur the expense, depending on requirements, of estimating the full state vector at the observation time points.

[22]This paper was reprinted twenty-five years later with an update as Moler and Van Loan (2003).

The evaluation of the likelihood in equation (29), in contrast, does not attempt an optimal prediction of x_{th} but rather models the time dependence between the η_{th} vectors parametrically. The computation of (29) is potentially troublesome as it involves the calculation of the determinant and inverse of the $nT \times nT$ covariance matrix Ω_η. However, the MA nature of η_{th} ensures that Ω_η is a sparse block-Toeplitz matrix with no more than $n(2p-1)$ non-zero elements in any row or column in the case of stocks and no more than $n(2p+1)$ non-zero elements in any row or column when flows are present. This sparsity can be exploited for computational advantages including speed and accuracy. Let M denote the $nT \times nT$ lower triangular Cholesky matrix with typical elements m_{ij} satisfying $MM' = \Omega_\eta$ and let $\xi = M^{-1}\eta$ with typical element ξ_i. Then $\eta'\Omega_\eta^{-1}\eta = \xi'\xi$ and $|\Omega_\eta| = |MM'| = |M|^2$ so that the log-likelihood can be written

$$\begin{aligned} \log L(\theta) &= -\frac{nT}{2}\log 2\pi - \frac{1}{2}\log|M|^2 - \frac{1}{2}\xi'\xi \\ &= -\frac{nT}{2}\log 2\pi - \sum_{i=1}^{nT}\log m_{ii} - \frac{1}{2}\sum_{i=1}^{nT}\xi_i^2, \end{aligned} \tag{31}$$

which follows because $|M| = \prod_{i=1}^{nT} m_{ii}$. Bergstrom (1983) showed that the elements of ξ can be computed recursively from the system $M\xi = \eta$ while Bergstrom (1990, chapter 7) showed that the elements of M converge rapidly to fixed limits as computations proceed within the matrix, resulting in savings in computational storage requirements.

One of the important features of a continuous time model is that the form of an exact discrete time representation is invariant to the sampling frequency of the observations. We are able to illustrate this aspect in the context of a small simulation exercise using a simple first-order stochastic differential equation in a scalar random variable $x(t)$, given by

$$dx(t) = ax(t)dt + \zeta(dt), \quad t > 0, \tag{32}$$

where we take $x(0) = 0$ for convenience and $E[\zeta(dt)^2] = \sigma^2 dt$. Assuming $x(t)$ to be a stock variable, suppose that the sequence $x_0, x_h, x_{2h}, \ldots, x_{Th}$ is observed, where h denotes the sampling interval and T is the number of discrete time observations. Then the exact discrete time model satisfied by the sequence of observations is a discrete time AR(1), regardless of the sampling interval; it is given by

$$x_{th} = f_h x_{th-h} + \eta_{th}, \quad t = 1, \ldots, N, \tag{33}$$

where $f_h = e^{ah}$ and η_{th} is white noise with variance $\sigma_\eta^2 = \sigma^2(e^{2ah} - 1)/(2a)$. Acknowledging that x_{th} is subject to temporal aggregation means that we focus on estimating a and σ^2 regardless of the sampling interval. However, ignoring this feature means that f_h would be estimated directly and estimates would suggest differing degrees of serial correlation depending on the value of h. Associated patterns of variation would also be observed in estimates of σ_η^2 owing to its dependence on h.

In order to assess these features we consider values of $h \in \{1/12, 1/6, 1/4, 1/3, 1/2, 1\}$ and $a \in \{-2, -1, -0.5, -0.1\}$ with $\sigma^2 = 1$. A total of 100,000 replications of each parameter combination were conducted, and we set the data span equal to $N = Th = 100$; this is the number of observations when $h = 1$. As the sampling interval falls the number of observations, $T = N/h$, rises, to a maximum of 1200 when $h = 1/12$. The data are generated at this highest frequency ($h = 1/12$) and then the lower frequency observations are selected, so that, for example, the observations for $h = 1/6$ correspond to every second observation in the $h = 1/12$ sequence, while those for $h = 1$ correspond to every twelfth observation. The maximum likelihood estimator of a can be shown to be equal to

$$\hat{a}_{ML} = \frac{1}{h} \log \hat{f}_h,$$

where $\hat{f}_h$ denotes the ordinary least squares (OLS) estimator of f_h in the autoregression (33). Clearly this is only feasible if $\hat{f}_h > 0$, and it is only for smaller values of f_h and T that it becomes a problem. In fact, the only cases where $\hat{f}_h < 0$ were for $a = -1$ and $a = -2$ when $T = 100$, where the proportions of replications affected were 0.00016 and 0.091, respectively. In these cases the estimates were removed and the summary statistics were computed with the remainder of the replications. In view of the results of Wang, Phillips and Yu (2011) we also compute an estimate of a based on the Euler approximation given by

$$(x_{th} - x_{th-h}) = ahx_{th-h} + u_{th},$$

where u_{th} is a serially uncorrelated random disturbance with variance $\sigma^2 h$. We denote the estimate of a obtained using this approximation by $\hat{a}_E$.

The results appear in Table 1 in which, for each value of a, the mean values and standard errors (across the replications) of $\hat{a}_{ML}$ and $\hat{a}_E$ are reported, as well as the means and standard errors of $\hat{f}_h$. In the latter case we also report the actual values of f_h. It can be seen clearly from Table 1 that the estimates of a using $\hat{a}_{ML}$, although slightly biased as expected, are all stable across the range of values of h, and although $\hat{a}_E$ has smaller bias than $\hat{a}_{ML}$ when $a = -0.1$ its performance in terms of bias deteriorates as a becomes more negative. This is in accordance with the results of Wang, Phillips and Yu (2011). It can also be seen that $\hat{a}_E$ has a smaller standard error than $\hat{a}_{ML}$ in all cases. The estimates of the discrete time autoregressive parameter f_h using $\hat{f}_h$ can be seen to depend clearly on the value of h. Although $\hat{f}_h$ is a reasonably good estimator of f_h the implications for the dependence properties of the variable x depend very much on the sampling interval chosen; the same is not true when the temporal aggregation is taken into account.

4. Empirical applications

There have been many applications of the methods of this chapter, most notably in the area of macroeconomic modelling to which, because it drove much of the early work, we devote section 4.3 below. Representative papers include, in the areas of asset allocation, Campbell, Chacko, Rodriguez and Viceira (2004); consumers' demand, Bergstrom and Chambers (1990) and Chambers (1992); uncovered interest parity, Diez de Los Rios and Sentana (2006); exchange rates, Renault, Sekkat and Szafarz (1998); short-term interest rate models, Nowman (1997), Yu and Phillips (2001), Phillips and Yu (2011); and empirical finance in general, Thornton and Chambers (2016). It is quite clear that, perhaps now more than ever, economic activity occurs continuously around the clock and yet, such is the undertaking required to measure this activity, published statistics cannot hope to provide real time information. Here we introduce two applications within macroeconomics which are illustrative of the use of an exact discrete representation to resolve this tension. One involves a univariate time series, namely consumer price inflation in the United Kingdom, while the other explores the important relationship between gross domestic product (GDP) in the United States of America and crude oil prices. As mentioned in section 2, the impact of time aggregation is to induce serial correlation in the disturbances, η_t, and so the ability of a continuous time specification to explain the observed serial correlation adequately is an important test of its suitability. In order to address this issue Bergstrom (1990, chapter 7) proposed a portmanteau-type test statistic based on the $n \times 1$ vectors of standard normal variates ξ_{th} $(t = 1, \ldots, T)$. Bergstrom's statistic is defined by

$$S_l = \frac{1}{n(T-l)} \sum_{r=1}^{l} \left(\sum_{t=l+1}^{T} \xi'_{th} \xi_{th-rh} \right)^2 ,$$

which has an approximate chi-squared distribution with l degrees of freedom (the number of lags used) under the null hypothesis that the model is correctly specified. For robustness we also report the Schwartz Bayesian model selection criterion (SBC) for each model. Each of these models is deliberately narrow in their focus. Modern economies are, of course, large, complex and inter-connected systems and so we finish with an overview of some of the large scale macroeconomic modelling carried out in continuous time.

4.1 Consumer prices and inflation

In a continuous time setting price inflation can be defined as the instantaneous rate of change of the price level i.e. $\pi(t) = D \log p(t)$. Consider the continuous time ARMA$(2,1)$ model for $\log p(t)$ given by

$$D^2 \log p(t) = \gamma_0 + A_1 D \log p(t) + A_0 \log p(t) + u(t) + \theta D u(t), \ t > 0, \tag{34}$$

where γ_0, A_1, A_0 and θ are scalars, and $u(t)$ is a mean zero uncorrelated process with variance σ_u^2. Under the condition that $A_0 = 0$ i.e. that $\log p(t)$ has a zero root in continuous time (and a unit root in discrete time), the implied law of motion for inflation becomes

$$D\pi(t) = \gamma_0 + A_1 \pi(t) + u(t) + \theta D u(t), \ t > 0. \tag{35}$$

Hence $\pi(t)$ satisfies a continuous time ARMA$(1,1)$ process which corresponds to a continuous time ARIMA$(2,1,1)$ process for $\log p(t)$.

Estimates of (34) with $A_0 = 0$ were obtained using monthly data for the UK consumer price index over the period January 1996 to March 2014, a total of 219 observations. The results are given in Table 2. The estimates of the parameters in the CARMA$(2,0)$ are well determined and there is no evidence of misspecification, at least as measured by Bergstrom's S_4 statistic. However, addition of the MA(1) component results in a statistically significant increase in the value of the maximised log-likelihood function – the likelihood ratio statistic for testing the null hypothesis that $\theta = 0$ is equal to 6.4634 with a marginal probability of 0.0110. The p-value of the S_4 statistic is far from significant, suggesting that the inclusion of the statistically significant MA(1) component yields an improved fit; this is also the inference drawn from a comparison of the SBC values for the two models.

4.2. Oil prices and the macroeconomy

Next, we explore the relationship between US output, as measured by real GDP in tens of billions of chained 2009 dollars, and the oil price, as measured by the price of West Texas Intermediate in dollars per barrel. The data are quarterly ranging from 1986 to 2013 quarter 3 from the Federal Reserve Bank of St Louis. In common with most authors who have examined these series, for example Hamilton (1996), we find that both processes show strong evidence of unit root behaviour, with augmented Dickey Fuller test statistics of -0.609 and -0.318 for GDP and the oil price respectively, but find no evidence of a cointegrating relationship. Non-stationary but non-cointegrated data are consistent with the specification in equation (15) with $A_0 = 0$. We define the 2×1 vector $x(t) = [GDP(t), Oil(t)]'$.

We consider two candidate models nested within a continuous time ARMA(2 ,1) model, the continuous time ARIMA(1,1,0) which has $A_0 = 0$, $p = 2$ and $q = 0$ and the continuous time CARIMA$(1, 1, 1)$, which has $q = 1$. The exact discrete representation of both models is an ARIMA$(1, 1, 1)$,

$$\Delta_h x_{th} = f_0 + F_1 \Delta x_{th-h} + \eta_{th}, \quad t = 3, \ldots, T, \tag{36}$$

where $\Delta_h x_{th} = x_{th} - x_{th-h}$ as in (24), with the CARIMA$(1, 1, 1)$ offering more flexibility in modelling the autocovariance structure of the discrete time disturbance η_{th}.

Results for the two models are presented in Table 3. The CARIMA$(1, 1, 1)$ is preferred by the SBC and the likelihood ratio test fails to reject the CARIMA$(1, 1, 1)$ in favour of the CARIMA$(1, 1, 0)$, with a test statistic of 24.566. The moving average coefficients in the first column of Θ are significant, reflecting the impact of lagged shocks to GDP on both GDP and oil prices. Both models have values for the Bergstrom S_1 and S_4 statistic in the acceptable region.

The literature has focussed on the question of whether changes in the oil price lead to changes in GDP, reflected in the top right element of the matrix A_1. It is noticeable that in the CARIMA$(1, 1, 0)$, neither of the coefficients on the rate of change of GDP or on the oil price is significant in the equation determining the other variable. When a moving average error is introduced to capture more complicated dynamics, however, the t-ratio on the top right element of A_1 is -2.36, indicating that growth in oil prices slows GDP, while that on the bottom left is 1.975, suggesting that growth in GDP accelerates oil price growth.

As a by-product of the estimation, to aid comparison with other models, the intercept vector, f_0, and autoregressive matrix, F_1 of the exact discrete time representation are also reported. These reinforce the point that the CARIMA$(1, 1, 1)$ model predicts a stronger reaction from one series to lagged changes in the other.

4.3. Macroeconometric modelling

While many of the applications of continuous time models and methods occur today in the area of empirical finance, much of the literature's early development was driven through the desire to make advances in the area of large-scale macroeconometric modelling. This was in no small part due to Rex Bergstrom who, in collaboration with a student, Clifford Wymer, produced in Bergstrom and Wymer (1976) the first continuous time macroeconometric model, the formulation and estimation of which represents one of the landmarks in the development of modern econometrics. The economy-wide model, which comprised thirteen equations (ten structural equations and three identities) in thirty-five key parameters, served as a prototype for later developments in macroeconomic modelling and the modelling of financial and commodity markets. It had innovative features beyond simply being formulated in continuous time: it was formulated as a dynamic disequilibrium model[23] involving a system of partial adjustment equations in the form of continuous time error-correction equations, where each causally dependent variable continually adjusts in response to the deviation from its partial equilibrium level; it embodied the intensive use of economic theory and other *a priori* information to support a parsimonious representation in the model parameters; and its design facilitated an analysis of its steady state and stability properties using methods developed earlier by Bergstrom (1967).[24]

An earlier comprehensive survey of continuous time macroeconomic modelling can be found in Bergstrom (1996), which includes the various stages of the Italian continuous time model of Gandolfo and Padoan (1984, 1990),[25] the economy-wide models contained in the volume edited by Gandolfo (1993) and, not least, the model by Bergstrom, Nowman and Wymer (1992) which signified the next stage of development. This model was the first to incorporate the exact methods that are the focus of this chapter; it used more realistic, second-order partial adjustment equations using the method for higher-order systems pioneered by Bergstrom (1983) and described in section 2.1 above; and, unlike the Bergstrom and Wymer (1976) model, incorporated exogenous variables. Estimating this model, which comprised fourteen equations with sixty-three parameters and eleven exogenous variables, required around a day's computing time on a CRAY X-MP/48 supercomputer, which at the time represented the cutting edge of computer technology. The Italian model was further developed into a system including non-linear equations by Gandolfo, Padoan, De Arcangelis and Wymer (1996), although estimation was facilitated through a linear approximation about sample means; see Wymer (1993) for details of the underlying estimation method. Wymer (1993, 1997, 2012) has developed a direct, full-information maximum likelihood approach to the estimation of such non-linear systems although, given the development of this literature, the properties of this estimator must currently be inferred from those of the estimator based on a linear approximation about sample means. Starting values for the procedure are readily obtained from applying the method of maximum likelihood to this linear approximation.

The theoretical basis for what could be seen as a third-stage continuous time model was provided by Bergstrom (1997), where unobservable stochastic trends are incorporated within the system of stochastic differential equations to take advantage of insights gained from the development of unit root econometrics that occurred in the discrete-time literature. The project was finally brought to fruition in a book by Bergstrom and Nowman (2007) that was published after Bergstrom's death in 2005. The model comprised a system of eighteen mixed first- and second-order non-linear differential equations with sixty-three structural parameters, thirty-three long-run parameters, twenty-seven speed-of-adjustment parameters and three drift parameters. Its linearisation about sample means results in precisely the

[23]See Hillinger (1996) for a discussion of the history and conceptual foundations of such models in macroeconometric modelling and Wymer (1996) for a similar discussion that focuses on continuous time.

[24]See Gandolfo (1981) for a textbook treatment.

[25]At the time of writing (July 2017), Pier Carlo Padoan is Italy's Minister of Economy and Finance, a position he has held since February 2014.

model considered by Bergstrom (1997). The parameter estimates and speed of adjustment parameters were all plausible and the model was seen, through an examination of its steady-state and stability properties, to generate plausible long-run behaviour. Its post-sample forecasting performance also compared favourably with a second-order VAR model with exogenous variables. The book provided a retrospect of what Rex Bergstrom achieved over a lifetime of research in the area of continuous time econometrics; a brief survey of this contribution with an emphasis on macroeconomic modelling is provided by Nowman (2009).

5. Concluding comments

This chapter has aimed to provide a survey of methods of continuous time modelling based on an exact discrete time representation. Such an approach is synonymous with the name of Rex Bergstrom whose pioneering contributions were instrumental in attracting the current authors to the field. Our survey has attempted to highlight the techniques involved with the derivation of an exact discrete time representation of an underlying continuous time model, providing specific details for a second-order linear system of stochastic differential equations. Issues of parameter identification, causality, nonstationarity, and mixed frequency data have also been addressed, all of which are important to consider in applications in economics and other disciplines. Although our focus has been on Gaussian estimation of the exact discrete time model we have also discussed alternative time domain (state space) and frequency domain approaches. Computational issues have also been explored, where here the focus is on the exploitation of sparse matrices and the computation of the matrix exponential. Two new empirical applications have been included along with a discussion of applications in the field of macroeconometric modelling. While our focus is, of necessity, oriented towards economics and econometrics, we hope that the material contained in this chapter will be of interest in the social and behavioural sciences more widely.

References

Aït-Sahalia, Y., 2007. Estimating continuous time models using discretely sampled data. In R. Blundell, T. Persson and W.K. Newey (eds.), *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society*, Cambridge University Press.

Aït-Sahalia, Y., Jacod, J., 2014. *High-Frequency Financial Econometrics.* Princeton University Press.

Bergstrom, A.R., 1966. Non-recursive models as discrete approximations to systems of stochastic differential equations. *Econometrica* 34, 173–182.

Bergstrom, A.R., 1967. *The Construction and Use of Economic Models.* English Universities Press.

Bergstrom, A.R., 1983. Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica* 51, 117–152.

Bergstrom, A.R., 1984. Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics, Vol. 2*, pp. 1145–1212, North-Holland.

Bergstrom, A.R., 1985. The estimation of parameters in nonstationary higher-order continuous-time dynamic models. *Econometric Theory* 1, 369–385.

Bergstrom, A.R., 1986. The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory* 2, 350–373.

Bergstrom, A.R., 1990. *Continuous Time Econometric Modelling.* Oxford University Press, Oxford.

Bergstrom, A.R., 1996. Survey of continuous time econometrics. In W.A. Barnett, G. Gandolfo and C. Hillinger (eds.), *Dynamic Disequilibrium Modeling*, pp. 3–25, Cambridge University Press.

Bergstrom, A.R., 1997. Gaussian estimation of mixed-order continuous-time dynamic models with unobservable stochastic trends from mixed stock and flow data. *Econometric Theory* 13, 467–505.

Bergstrom, A.R., Chambers, M.J., 1990. Gaussian estimation of a continuous time model of demand for consumer durable goods with applications to demand in the United Kingdom, 1973–1984. In A.R. Bergstrom (ed.), *Continuous Time Econometric Modelling*, pp.279–319, Oxford University Press, Oxford.

Bergstrom, A.R., Nowman, K.B., 2007. *A Continuous Time Econometric Model of the United Kingdom with Stochastic Trends.* Cambridge University Press, Cambridge.

Bergstrom, A.R., Nowman, K.B., Wymer, C.R., 1992. Gaussian estimation of a second order continuous time macroeconometric model of the United Kingdom. *Economic Modelling* 9, 313–351.

Bergstrom, A.R., Wymer, C.R., 1976. A model of disequilibrium neoclassical growth and its application to the United Kingdom. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 267–327, North-Holland.

Blevins, J.R., 2017. Identifying restrictions for finite parameter continuous time models with discrete time data. *Econometric Theory* 33, 739–754.

Breitung, J., Swanson, N.R., 2002. Temporal aggregation and spurious instantaneous causality in multiple time series models. *Journal of Time Series Analysis* 23, 651–665.

Brockwell, P.J., 2014. Recent results in the theory and applications of CARMA processes. *Annals of the Institute of Statistical Mathematics* 66, 647–685.

Campbell, J.Y., Chacko, G., Rodriguez, J., Viceira, L.M., 2004. Strategic asset allocation in a continuous-time VAR model. *Journal of Economic Dynamics and Control* 28, 2195–2214.

Chambers, M.J., 1990. *Durability and Consumers' Demand: Gaussian Estimation of Some Continuous Time Models.* PhD Thesis, University of Essex, Colchester.

Chambers, M.J., 1992. Estimation of a continuous time dynamic demand system. *Journal of Applied Econometrics* 7, 53–64.

Chambers, M.J., 1999. Discrete time representation of stationary and non-stationary continuous time systems. *Journal of Economic Dynamics and Control* 23, 619–639.

Chambers, M.J., 2001. Temporal aggregation and the finite sample performance of spectral regression estimators in cointegrated systems: a simulation study. *Econometric Theory* 17, 591–607.

Chambers, M.J., 2009. Discrete time representations of cointegrated continuous time models with mixed sample data. *Econometric Theory* 25, 1030–1049.

Chambers, M.J., 2011. Cointegration and sampling frequency. *Econometrics Journal* 14, 156–185.

Chambers, M.J., 2013. Jackknife estimation of stationary autoregressive models. *Journal of Econometrics* 172, 142–157.

Chambers, M.J., 2016. The estimation of continuous time models with mixed frequency data. *Journal of Econometrics* 193, 390–404.

Chambers, M.J., Kyriacou, M., 2013. Jackknife estimation with a unit root. *Statistics and Probability Letters* 83, 1677–1682.

Chambers, M.J., McCrorie, J.R., 2006. Identification and estimation of exchange rate models with unobservable fundamentals. *International Economic Review* 47, 573–582.

Chambers, M.J., McCrorie, J.R., 2007. Frequency domain estimation of temporally aggregated Gaussian cointegrated systems. *Journal of Econometrics* 136, 1–29.

Chambers, M.J., Thornton, M.A., 2012. Discrete time representations of continuous time ARMA processs. *Econometric Theory* 28, 219–238.

Christiano, L.J., Eichenbaum, M.S., 1987. Temporal aggregation and structural inference in macroeconomics. *Carnegie-Rochester Conference Series on Public Policy* 26, 63–130.

Comte, F., 1999. Discrete and continuous time cointegration. *Journal of Econometrics* 88, 207–226.

Comte, F., Renault, E., 1996. Noncausality in continuous time models. *Econometric Theory* 12, 215-256.

Corradi, V., 1997. Comovements between diffusion processes: characterization, estimation and testing. *Econometric Theory* 13, 646–666.

Culver, W., 1966. On the existence and uniqueness of the real logarithm of a matrix. *Proceedings of the American Mathematical Society* 16, 1146–1151.

Diez de los Rios, A., Sentana, E., 2011. Testing uncovered interest parity: a continuous-time approach. *International Economic Review* 52, 1215–1251.

Engle, R.F., Granger, C.W.J., 1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* 55, 251-276.

Fergusson, K., Platen, E., 2015. Application of maximum likelihood estimation to stochastic short rate models. *Annals of Financial Economics* 10, 1–26.

Florens, J.-P., Fougère, D., 1996. Noncausality in continuous time. *Econometrica* 64, 1195–1212.

Gandolfo, G., 1981. *Quantitative Analysis and Econometric Estimation of Continuous Time Models*. North-Holland.

Gandolfo, G., 1993. *Continuous Time Econometrics: Theory and Applications*. Chapman and Hall.

Gandolfo, G., Padoan, P.C., 1984. *A Disequilibrium Model of the Real and Financial Accumulation in an Open Economy*. Springer Verlag.

Gandolfo, G., Padoan, P.C., 1990. The Italian continuous time model, theory and empirical results. *Economic Modelling* 7, 91–132.

Gandolfo, G., Padoan, P.C., De Arcangelis, G., Wymer, C.R., 1996. Nonlinear estimation of a nonlinear continuous time model. In W.A. Barnett, G. Gandolfo and C. Hillinger (eds.), *Dynamic Disequilibrium Modeling*, pp. 127–150, Cambridge University Press.

Granger, C.W.J., 1966. The typical spectral shape of an economic variable. *Econometrica* 34, 150–161.

Hamerle, A., Nagl, W., Singer, H., 1991. Problems with the estimation of stochastic differential equations using structural equations models. *Journal of Mathematical Sociology* 16, 201–220.

Hamerle, A., Singer, H., Nagl, W., 1993. Identification and estimation of continuous time dynamic systems with exogenous variables using panel data. *Econometric Theory* 9, 283–295.

Hamilton, J.D., 1996. This is what happened to the oil price-macroeconomy relationship. *Journal of Monetary Economics* 38, 215–220.

Hansen L.P., Sargent, T.J., 1983. The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica* 51, 377–387.

Hansen L.P., Sargent, T.J., 1991. Identification of continuous time rational expectations from discrete data. In L.P. Hansen and T.J. Sargent (eds.), *Rational Expectations Econometrics*, pp. 219–235, Westview Press.

Harvey, A.C., 1989. *Forecasting Structual Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Harvey, A.C., Stock, J.H., 1985. The estimation of higher-order continuous time autoregressive models. *Econometric Theory* 1, 97–117.

Harvey, A.C., Stock, J.H., 1988. Continuous time autoregressive models with common stochastic trends. *Journal of Economic Dynamics and Control* 12, 365–384.

Harvey, A.C., Stock, J.H., 1989. Estimating integrated higher-order continuous time autoregressions with an application to money-income causality. *Journal of Econometrics* 42, 313–336.

Higham, N.J., 2009. The scaling and squaring method for the matrix exponential revisited. *SIAM Review* 51, 747–764.

Higham, N.J., Al-Mohy, A.H., 2010. Computing matrix functions. *Acta Numerica* 19, 159–208.

Hillinger, C., 1996. Dynamic disequilibrium economics: history, conceptual foundations, possible futures. In W.A. Barnett, G. Gandolfo and C. Hillinger (eds.), *Dynamic Disequilibrium Modeling*, pp. 27–65, Cambridge University Press.

Jewitt, G., McCrorie, J.R., 2005. Computing estimates of continuous time macroeconometric models on the basis of discrete data. *Computational Statistics and Data Analysis* 49, 397–416.

Jones, R.H., 1981. Fitting a continuous time autoregression to discrete data. In D.F. Findlay (ed.), *Applied Time Series Analysis II*, pp. 651–682. New York: Academic Press.

Marcellino, M., 1999. Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics* 17, 129–136.

McCrorie, J.R., 1996. *Some Topics in the Estimation of Continuous Time Econometric Models*. PhD Thesis, University of Essex, Colchester.

McCrorie, J.R., 2000. Deriving the exact discrete analog of a continuous time system. *Econometric Theory* 16, 998–1015.

McCrorie, J.R., 2001. Interpolating exogenous variables in continuous time dynamic models. *Journal of Economic Dynamics and Control* 25, 1399–1427.

McCrorie, J.R., 2003. The problem of aliasing in identifying finite parameter continuous time models. *Acta Applicandae Mathematicae* 79, 9–16.

McCrorie, J.R., 2009. Estimating continuous-time models on the basis of discrete data via an exact discrete analog. *Econometric Theory* 25, 1120–1137.

McCrorie, J.R., Chambers, M.J., 2006. Granger causality and the sampling of economic processes. *Journal of Econometrics* 132, 311–336.

Moler, C.B., Van Loan, C.F., 1978. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review* 20, 801–836.

Moler, C.B., Van Loan, C.F., 2003. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review* 45, 3–49.

Nowman, K.B., 1997. Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 52, 1695–1703.

Nowman, K.B., 2009. Albert Rex Bergstrom's contributions to continuous time macroeconometric modeling. *Econometric Theory* 25, 1087–1098.

Oud, J., Singer, H., 2008. Continuous time modelling of panel data: SEM versus filter techniques. *Statistica Neerlandica* 62, 4–28.

Phillips, P.C.B., 1972. The structural estimation of a stochastic differential equation system. *Econometrica* 40, 1021–1041.

Phillips, P.C.B., 1973. The problem of identification in finite parameter continuous time models. *Journal of Econometrics* 1, 351–362.

Phillips, P.C.B., 1974. The treatment of flow data in the estimation of continuous time systems. Paper presented at the European Meeting of the Econometric Society, Grenoble, 1974; available as Discussion Paper No. 60, University of Essex, Department of Economics, 1975. An abridged version was published as Chapter 15 in A.R. Bergstrom, A.J.L. Catt, M.H. Peston and B.D.J. Silverstone (eds.), *Stability and Inflation: Essays in Memory of A.W.H. Phillips*, pp. 257–274, John Wiley and Sons, New York, 1978.

Phillips, P.C.B., 1987. Time series regression with a unit root. *Econometrica* 55, 277–301.

Phillips, P.C.B., 1991. Error correction and long-run equilibrium in continuous time. *Econometrica* 59, 967–980.

Phillips, P.C.B., Yu, J., 2005. Jackknifing bond prices. *Review of Financial Studies* 18, 707–742.

Phillips, P.C.B., Yu, J., 2009. Maximum likelihood and Gaussian estimation of continuous time models in Finance. In T.G. Anderson, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.), *Handbook of Financial Time Series*, pp. 707–742, Springer-Verlag.

Phillips, P.C.B., Yu, J., 2011. Corrigendum to "A Gaussian approach for continuous time models of the short term interest rate." *Econometrics Journal* 14, 126–129.

Renault, E., Sekkat, K., Szafarz, A., 1998. Testing for spurious causality in exchange rates. *Journal of Empirical Finance* 5, 47–66.

Robinson, P.M., 1976. Fourier estimation of continuous time models. In A.R. Bergstrom (ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 215–266. North-Holland, Amsterdam.

Robinson, P.M., 1992. Review of A.R. Bergstrom's 'Continuous Time Econometric Modelling.' *Econometric Theory* 8, 571–579.1

Robinson, P.M., 1993. Continuous-time models in econometrics: closed and open systems, stocks and flows. In P.C.B. Phillips (ed.), *Models, Methods, and Applications of Econometrics: Essays in Honor of A.R. Bergstrom*, pp. 71–90. Blackwell, Oxford.

Rozanov, Y.A., 1967. *Stationary Random Processes.* Holden-Day, San Francisco.

Singer, H., 1995. Analytical score function for irregularly sampled continuous time stochastic processes with control variables and missing values. *Econometric Theory* 11, 721–735.

Singer, H., 2007. Stochastic differential equation models with sampled data. In K. van Montfort, J. Oud and A. Satorra (eds.), *Longitudinal Models in the Behavioural and Related Sciences*, pp. 73–106. Lawrence Erlbaum Associates.

Stock, J.H., 1987. Temporal aggregation and structural inference in macroeconomics: a comment. *Carnegie-Rochester Conference Series on Public Policy* 26, 131–140.

Thornton, M.A., 2009. *Information and Aggregation: The Econometrics of Dynamic Models of Consumption under Cross-Sectional and Temporal Aggregation.* PhD Thesis, University of Essex, Colchester.

Thornton, M.A., Chambers, M.J., 2013. Temporal aggregation in macroeconomics. In N. Hashimzade and M.A. Thornton (eds), *Handbook of Research Methods and Applications in Empirical Macroeconomics*, pp. 289–309, Edward Elgar.

Thornton, M.A., Chambers, M.J., 2016. The exact discretisation of CARMA models with applications in finance. *Journal of Empirical Finance* 38, 739–761.

Thornton, M.A., Chambers, M.J., 2017. Continuous time ARMA processes: discrete time representation and likelihood evaluation. *Journal of Economic Dynamics and Control* 79, 48–65.

Van Loan, C.F., 1978. Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control* 23, 395–404.

Wang, X., Phillips, P.C.B., Yu, J., 2011. Bias in estimating multivariate and univariate diffusions. *Journal of Econometrics* 161, 228–245.

Wymer, C.R., 1972. Econometric estimation of stochastic differential equations. *Econometrica* 40, 565–577.

Wymer, C.R., 1976. Continuous time models in macroeconomics: specification and estimation. Paper prepared for the SSRC Ford Foundation Conference on "Macroeconomic Policy and Adjustment in Open Economies" at Fanhams Hall, Ware, England, April 28-May 1, 1976. Published as Section I of Chapter 3 in G. Gandolfo (ed.), *Continuous Time Econometrics: Theory and Applications*, pp. 35–59, Chapman and Hall, London, 1993.

Wymer, C.R., 1993. Estimation of nonlinear continuous-time models from discrete data. In P.C.B. Phillips (ed.), *Models, Methods and Applications of Econometrics: Essays in Honor of A.R. Bergstrom*, pp. 91–114, Blackwell, Oxford.

Wymer, C.R., 1996. The role of continuous time disequilibrium models in macroeconomics. In W.A. Barnett, G. Gandolfo and C. Hillinger (eds.), *Dynamic Disequilibrium Modeling*, pp. 67–123, Cambridge University Press.

Wymer, C.R., 1997. Structural non-linear continuous-time models in econometrics. *Macroeconomic Dynamics* 1, 518–548.

Wymer, C.R., 2012. Continuous time econometrics of structural models. *Studies in Nonlinear Dynamics and Econometrics* 16, 1–28.

Yu, J., 2012. Bias in the estimation of the mean reversion parameter in continuous time models. *Journal of Econometrics* 169, 114–122.

Yu, J., 2014. Econometric analysis of continuous time models: a survey of Peter Phillipss work and some new results. *Econometric Theory* 30, 737–774.

Yu, J., Phillips, P.C.B., 2001. A Gaussian approach for continuous time models of the short-term interest rate. *Econometrics Journal* 4, 221–225.

Zadrozny, P., 1988. Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies. *Econometric Theory* 4, 108–124.

Appendix

The Gauss code below was used in the simulation exercise. Note that n is used in the code as the data span and t is the sample size, whereas in the text in section 3 it is T and N, respectively, that are used for these quantities.

```
/* Simulation of continuous time AR(1) process at different frequencies */

new;

a=-1.0;                    /* Continuous time AR parameter        */
hv=1|1/2|1/3|1/4|1/6|1/12; /* Discrete time sampling intervals    */
n=100;                     /* Data span                           */
x0=0;                      /* Initial value                       */
nreps=100000;              /* Number of replications              */
s2=1;                      /* Continuous time innovation variance */
rndseed 6665;              /* seed for random numbers             */

rhv=rows(hv);
tv=n./hv;
maxt=maxc(tv);
hmin=minc(hv);
hrel=hv/hmin;
eahm=exp(a*hmin);
e2ahm=exp(2*a*hmin);
eahv=exp(a*hv);
s2m=s2*(e2ahm-1)/(2*a);
sm=sqrt(s2m);
cta=zeros(nreps,rhv);      /* nreps times number of h values      */
dta=cta; eta=cta; nogood=0;

for i (1,nreps,1);
    u=sm*rndn(maxt,1);
    xm=datagen(u);         /* maxt times 1                        */
    for hi (1,rhv,1);
        h=hv[hi,1];
        t=tv[hi,1];
        xh=reshape(xm,tv[hi,1],hrel[hi,1]);
        x=xh[.,hrel[hi,1]];
        bhat=x[2:t,1]/x[1:t-1,1];
        if bhat le 0; ahat=0; nogood=nogood+1;
        else; ahat=ln(bhat)/h;
        endif;
        ehat=(x[2:t,1]-x[1:t-1,1])/(h*x[1:t-1,1]);
        cta[i,hi]=ahat;
        dta[i,hi]=bhat;
        eta[i,hi]=ehat;
    endfor;
endfor;

stop;

proc datagen(e);
    local x;
    x = recserar(e, x0, eahm);
    retp( x );
endp;
```

Table 1. Simulation results: means and standard errors of estimators

h	$\hat{a}_{ML}$	$\hat{a}_E$	$\hat{f}_h$	f_h	$\hat{a}_{ML}$	$\hat{a}_E$	$\hat{f}_h$	f_h
		$a = -0.1$				$a = -0.5$		
1	−0.1213	−0.1127	0.8873	0.9048	−0.5292	−0.4050	0.5950	0.6065
	(0.0591)	(0.0504)	(0.0504)		(0.1452)	(0.0814)	(0.0814)	
1/2	−0.1206	−0.1163	0.9418	0.9512	−0.5235	−0.4578	0.7711	0.7788
	(0.0567)	(0.0524)	(0.0262)		(0.1217)	(0.0909)	(0.0460)	
1/3	−0.1203	−0.1174	0.9609	0.9672	−0.5219	−0.4772	0.8409	0.8465
	(0.0559)	(0.0530)	(0.0177)		(0.1153)	(0.0957)	(0.0319)	
1/4	−0.1202	−0.1181	0.9705	0.9753	−0.5214	−0.4875	0.8781	0.8825
	(0.0555)	(0.0534)	(0.0133)		(0.1125)	(0.0979)	(0.0245)	
1/6	−0.1201	−0.1186	0.9802	0.9835	−0.5207	−0.4979	0.9170	0.9200
	(0.0551)	(0.0537)	(0.0090)		(0.1097)	(0.1000)	(0.0167)	
1/12	−0.1200	−0.1193	0.9901	0.9917	−0.5203	−0.5088	0.9576	0.9592
	(0.0548)	(0.0541)	(0.0045)		(0.1072)	(0.1024)	(0.0085)	
		$a = -1.0$				$a = -2.0$		
1	−1.0581	−0.6388	0.3612	0.3679	−2.1249	−0.8667	0.1333	0.1353
	(0.3030)	(0.0934)	(0.0934)		(0.8435)	(0.0988)	(0.0988)	
1/2	−1.0291	−0.7989	0.6006	0.6065	−2.0554	−1.2716	0.3642	0.3679
	(0.1951)	(0.1140)	(0.0570)		(0.3853)	(0.1318)	(0.0659)	
1/3	−1.0247	−0.8644	0.7119	0.7165	−2.0343	−1.4698	0.5101	0.5134
	(0.1746)	(0.1225)	(0.0408)		(0.2986)	(0.1488)	(0.0496)	
1/4	−1.0234	−0.9003	0.7749	0.7788	−2.0291	−1.5861	0.6035	0.6065
	(0.1662)	(0.1275)	(0.0319)		(0.2682)	(0.1597)	(0.0399)	
1/6	−1.0217	−0.9377	0.8437	0.8465	−2.0244	−1.7148	0.7142	0.7165
	(0.1583)	(0.1327)	(0.0221)		(0.2420)	(0.1716)	(0.0286)	
1/12	−1.0208	−0.9777	0.9185	0.9200	−2.0219	−1.8590	0.8451	0.8465
	(0.1513)	(0.1386)	(0.0115)		(0.2205)	(0.1858)	(0.0155)	

Table 2. Estimates for Inflation

	CARMA$(2,0)$	CARMA$(2,1)$
γ_0	0.0261	0.0013
	(0.0038)	(0.0006)
A_1	-14.8432	-0.7362
	(0.0046)	(0.3425)
θ	0.0000	-1.9916
		(0.9772)
σ_u	0.0562	0.0020
	(0.0027)	(0.0010)
$\log L$	909.8639	913.0956
SBC	-1803.5607	-1804.6349
S_4	0.4427	0.9942

Standard errors in parentheses; entries for S_4 are p-values.

Table 3. Estimates of CARIMA models for GDP and oil prices

	CARIMA(1, 1, 0)			CARIMA(1, 1, 1)				
	a_0	A_1		a_0	A_1		Θ	
Continuous time model parameters								
GDP	3.9720	−0.4966	−0.3474	4.1819	−0.5107	−0.4729	−0.3648	−0.9745
	(4.3642)	(0.2368)	(0.4838)	(4.3192)	(0.1890)	(0.2004)	(0.1829)	(1.2448)
Oil price	2.8840	−0.2382	−1.2451	−5.5287	1.0499	−2.4061	0.3253	−0.7239
	(1.7099)	(0.6098)	(0.6241)	(1.7247)	(0.5317)	(0.6380)	(0.1594)	(0.7642)
$\log L$		−737.8751				−725.5962		
SBC		1518.1359				1512.4163		
S_1		0.4328				0.6619		
S_4		0.1875				0.1945		
	f_0	F_1		f_0	F_1			
Exact discrete time model parameters								
GDP	3.1898	0.6033	−0.2960	2.8670	0.6797	−0.5649		
Oil price	1.1864	−0.0784	0.3289	−0.7302	0.2376	−0.1179		

Standard errors in parentheses; entries for S_1 and S_4 are p-values.